Second order asymptotic efficiency for a Poisson process

Samvel B. GASPARYAN

Joint work with Yu.A. Kutoyants

Université du Maine, Le Mans, France

masiv6@gmail.com

10 June, 2016
Rennes

Outline

Inhomogeneous Poisson process

- In non-parametric estimation the unknown object is a function.

Inhomogeneous Poisson process

- In non-parametric estimation the unknown object is a function.
- We observe a periodic Poisson process with a known period τ

$$X^T = \{X(t),\ t \in [0, T]\},\ T = n\tau.$$

Inhomogeneous Poisson process

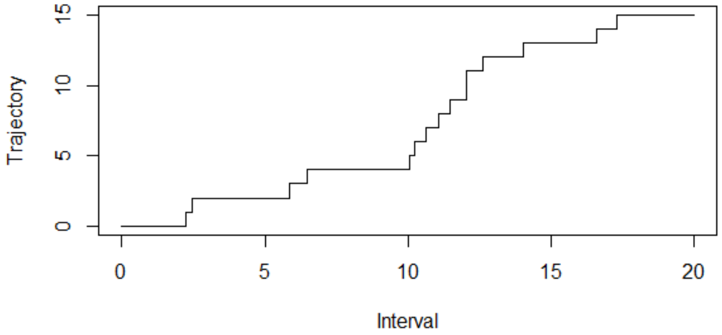
- In non-parametric estimation the unknown object is a function.
- We observe a periodic Poisson process with a known period τ

$$X^T = \{X(t),\ t \in [0, T]\},\ T = n\tau.$$

- $X(0) = 0$, has independent increments and there exists a positive, increasing function $\Lambda(t)$ s.t. for all $t \in [0, T]$

$$\mathbf{P}(X(t) = k) = \frac{[\Lambda(t)]^k}{k!} e^{-\Lambda(t)},\ k = 0, 1, \cdots .$$

Trajectory of a Poisson process



Models description

- In the definition

$$\mathbf{P}(X(t)=k)=\frac{[\Lambda(t)]^k}{k!}e^{-\Lambda(t)},\ k=0,1,\cdots.$$

Models description

- In the definition

$$\mathbf{P}(X(t) = k) = \frac{[\Lambda(t)]^k}{k!} e^{-\Lambda(t)}, \; k = 0, 1, \cdots .$$

- The function $\Lambda(\cdot)$ is called the mean function since $\mathbf{E}X(t) = \Lambda(t)$.

Models description

- In the definition

$$\mathbf{P}(X(t)=k)=\frac{[\Lambda(t)]^k}{k!}e^{-\Lambda(t)},\ k=0,1,\cdots.$$

- The function $\Lambda(\cdot)$ is called the mean function since $\mathbf{E}X(t)=\Lambda(t)$.
- We consider the case were $\Lambda(\cdot)$ is absolutely continuous $\Lambda(t)=\int_0^t\lambda(s)\mathrm{d}s$.

Models description

- In the definition

$$\mathbf{P}(X(t)=k)=\frac{[\Lambda(t)]^k}{k!}e^{-\Lambda(t)},\ k=0,1,\cdots.$$

- The function $\Lambda(\cdot)$ is called the mean function since $\mathbf{E}X(t)=\Lambda(t)$.
- We consider the case were $\Lambda(\cdot)$ is absolutely continuous $\Lambda(t)=\int_0^t\lambda(s)\mathrm{d}s$.
- The positive function $\lambda(\cdot)$ is called the intensity function and the periodicity of a Poisson process means the periodicity of its intensity function

$$\lambda(t)=\lambda(t+k\tau),\ t\in[0,\tau],\ k\in\mathcal{Z}_+.$$

Mean and the Intensity functions

- With the notations

$$X_j(t) = X((j-1)\tau + t) - X((j-1)\tau),\ t \in [0,\tau],$$
$$X_j = \{X_j(t),\ t \in [0,\tau]\},\ j = 1,\cdots,n,$$

Mean and the Intensity functions

- With the notations

$$X_j(t) = X((j-1)\tau + t) - X((j-1)\tau),\ t \in [0,\tau],$$
$$X_j = \{X_j(t),\ t \in [0,\tau]\},\ j = 1,\cdots,n,$$

- We get an i.i.d. model $X^n = (X_1, X_2, \cdots, X_n)$ generated from a Poisson process.

Mean and the Intensity functions

- With the notations

$$X_j(t) = X((j-1)\tau + t) - X((j-1)\tau),\ t \in [0,\tau],$$
$$X_j = \{X_j(t),\ t \in [0,\tau]\},\ j = 1,\cdots,n,$$

- We get an i.i.d. model $X^n = (X_1, X_2, \cdots, X_n)$ generated from a Poisson process.
- Estimation problems of $\lambda(t),\ t \in [0,\tau]$ and $\Lambda(t),\ t \in [0,\tau]$ are completely different.

Mean and the Intensity functions

- With the notations

$$X_j(t) = X((j-1)\tau + t) - X((j-1)\tau),\ t \in [0,\tau],$$
$$X_j = \{X_j(t),\ t \in [0,\tau]\},\ j = 1, \cdots, n,$$

- We get an i.i.d. model $X^n = (X_1, X_2, \cdots, X_n)$ generated from a Poisson process.
- Estimation problems of $\lambda(t),\ t \in [0,\tau]$ and $\Lambda(t),\ t \in [0,\tau]$ are completely different.
- We would like to have Hájek-Le Cam type lower bounds for function estimation

$$\lim_{\delta \downarrow 0} \varliminf_{n \to +\infty} \sup_{|\theta - \theta_0| \leq \delta} n\mathbf{E}_\theta(\bar{\theta}_n - \theta)^2 \geq \frac{1}{\mathrm{I}(\theta_0)}.$$

Mean function estimation

- We consider the estimation problem of the mean function $\{\Lambda(t),\ t \in [0,\tau]\}$.

Mean function estimation

- We consider the estimation problem of the mean function $\{\Lambda(t),\ t \in [0,\tau]\}$.
- Each measurable function $\bar{\Lambda}_n(t) = \bar{\Lambda}_n(t, X^n)$ of observations is an estimator for $\Lambda(t)$.

Mean function estimation

- We consider the estimation problem of the mean function $\{\Lambda(t),\ t \in [0,\tau]\}$.
- Each measurable function $\bar{\Lambda}_n(t) = \bar{\Lambda}_n(t, X^n)$ of observations is an estimator for $\Lambda(t)$.
- To assess the quality of an estimator we use the MISE

$$\mathbf{E}_\Lambda ||\bar{\Lambda}_n - \Lambda||^2 = \mathbf{E}_\Lambda \int_0^\tau (\bar{\Lambda}_n(t) - \Lambda(t))^2 \mathrm{d}t.$$

Mean function estimation

- We consider the estimation problem of the mean function $\{\Lambda(t),\ t \in [0,\tau]\}$.
- Each measurable function $\bar{\Lambda}_n(t) = \bar{\Lambda}_n(t, X^n)$ of observations is an estimator for $\Lambda(t)$.
- To assess the quality of an estimator we use the MISE

$$\mathbf{E}_\Lambda ||\bar{\Lambda}_n - \Lambda||^2 = \mathbf{E}_\Lambda \int_0^\tau (\bar{\Lambda}_n(t) - \Lambda(t))^2 \mathrm{d}t.$$

- The simplest estimator is the *empirical mean function*

$$\hat{\Lambda}_n(t) = \frac{1}{n}\sum_{j=1}^{n} X_j(t),\ t \in [0,\tau].$$

The basic equality for the EMF

- The following basic equality for the EMF implies two things

$$\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 = \int_0^\tau \Lambda(t)\mathrm{d}t.$$

The basic equality for the EMF

- The following basic equality for the EMF implies two things

$$\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n-\Lambda)||^2=\int_0^\tau \Lambda(t)\mathrm{d}t.$$

1. The EMF is consistent with the classical rate of convergence $\sqrt{n}$.

The basic equality for the EMF

- The following basic equality for the EMF implies two things

$$\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 = \int_0^\tau \Lambda(t)\mathrm{d}t.$$

1. The EMF is consistent with the classical rate of convergence $\sqrt{n}$.
2. The asymptotic variance (which is non-asymptotic) of the EMF is $\int_0^\tau \Lambda(t)\mathrm{d}t$.

The basic equality for the EMF

- The following basic equality for the EMF implies two things

$$\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 = \int_0^\tau \Lambda(t)\mathrm{d}t.$$

1. The EMF is consistent with the classical rate of convergence $\sqrt{n}$.
2. The asymptotic variance (which is non-asymptotic) of the EMF is $\int_0^\tau \Lambda(t)\mathrm{d}t$.

- Can we have better rate of convergence or smaller asymptotic variance for an estimator?

As. efficiency of the EMF

- The EMF is asymptotically efficient among all estimators $\bar{\Lambda}_n(t)$,

As. efficiency of the EMF

- The EMF is asymptotically efficient among all estimators $\bar{\Lambda}_n(t)$,
- Kutoyants' result-for all estimators $\bar{\Lambda}_n(t)$

$$\lim_{\delta\downarrow 0} \underline{\lim}_{n\to+\infty} \sup_{\Lambda\in V_\delta} \mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 \geq \int_0^\tau \Lambda^*(t)\mathrm{d}t,$$

with $V_\delta = \{\Lambda : \sup_{0\leq t\leq\tau} |\Lambda(t) - \Lambda^*(t)| \leq \delta\}, \delta > 0.$

As. efficiency of the EMF

- The EMF is asymptotically efficient among all estimators $\bar{\Lambda}_n(t)$,
- Kutoyants' result-for all estimators $\bar{\Lambda}_n(t)$

$$\lim_{\delta \downarrow 0} \varliminf_{n \to +\infty} \sup_{\Lambda \in V_\delta} \mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 \geq \int_0^\tau \Lambda^*(t) \mathrm{d}t,$$

with $V_\delta = \{\Lambda : \sup_{0 \leq t \leq \tau} |\Lambda(t) - \Lambda^*(t)| \leq \delta\}, \delta > 0.$

- Reformulation

$$\varliminf_{n \to +\infty} \sup_{\Lambda \in \mathcal{F}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t) \mathrm{d}t \right) \geq 0.$$

As. efficiency of the EMF

- The EMF is asymptotically efficient among all estimators $\bar{\Lambda}_n(t)$,
- Kutoyants' result-for all estimators $\bar{\Lambda}_n(t)$

$$\lim_{\delta \downarrow 0} \varliminf_{n \to +\infty} \sup_{\Lambda \in V_\delta} \mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 \geq \int_0^\tau \Lambda^*(t) \mathrm{d}t,$$

with $V_\delta = \{\Lambda : \sup_{0 \leq t \leq \tau} |\Lambda(t) - \Lambda^*(t)| \leq \delta\}, \delta > 0.$

- Reformulation

$$\varliminf_{n \to +\infty} \sup_{\Lambda \in \mathcal{F}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t) \mathrm{d}t \right) \geq 0.$$

- $\mathcal{F} \subset \mathrm{L}_2[0, \tau]$ is a sufficiently "rich", bounded set.

As. efficiency of the EMF

- The EMF is asymptotically efficient among all estimators $\bar{\Lambda}_n(t)$,
- Kutoyants' result-for all estimators $\bar{\Lambda}_n(t)$

$$\lim_{\delta \downarrow 0} \varliminf_{n \to +\infty} \sup_{\Lambda \in V_\delta} \mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 \geq \int_0^\tau \Lambda^*(t) \mathrm{d}t,$$

with $V_\delta = \{\Lambda : \sup_{0 \leq t \leq \tau} |\Lambda(t) - \Lambda^*(t)| \leq \delta\}$, $\delta > 0$.

- Reformulation

$$\varliminf_{n \to +\infty} \sup_{\Lambda \in \mathcal{F}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t) \mathrm{d}t \right) \geq 0.$$

- $\mathcal{F} \subset \mathrm{L}_2[0, \tau]$ is a sufficiently "rich", bounded set.
- Can we have other asymptotically efficient estimators?

Efficient estimators

- Existence of other as. efficient estimators depends on the regularity conditions imposed on unknown $\Lambda(\cdot)$.

Efficient estimators

- Existence of other as. efficient estimators depends on the regularity conditions imposed on unknown $\Lambda(\cdot)$.
- In other words, it depends on the choice of the set $\mathcal{F}$ in

$$\underline{\lim}_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t\right) \geq 0.$$

Efficient estimators

- Existence of other as. efficient estimators depends on the regularity conditions imposed on unknown $\Lambda(\cdot)$.
- In other words, it depends on the choice of the set $\mathcal{F}$ in

$$\underline{\lim}_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq 0.$$

- Demanding existence of derivatives of higher order of the unknown function, we can enlarge the class of as. efficient estimators.

First results

- At first, consider the L_2 ball with a center Λ^*

$$\mathcal{B}(R) = \{\Lambda : \ ||\Lambda - \Lambda^*||^2 \leq R, \ \Lambda^*(\tau) = \Lambda(\tau)\}.$$

First results

- At first, consider the L_2 ball with a center Λ^*

$$\mathcal{B}(R) = \{\Lambda : \ ||\Lambda - \Lambda^*||^2 \leq R, \, \Lambda^*(\tau) = \Lambda(\tau)\}.$$

- Consider a kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^\tau K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

First results

- At first, consider the L_2 ball with a center Λ^*

$$\mathcal{B}(R) = \{\Lambda : \, ||\Lambda - \Lambda^*||^2 \leq R, \, \Lambda^*(\tau) = \Lambda(\tau)\}.$$

- Consider a kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^\tau K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

- Kernels satisfy

$$K_n(u) \geq 0, \, u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right], \quad \int_{-\frac{\tau}{2}}^{-\frac{\tau}{2}} K_n(u)\mathrm{d}u = 1, \, n \in \mathcal{N},$$

and we continue them τ periodically on the whole real line $\mathbf{R}$

$$K_n(u) = K_n(-u), \quad K_n(u) = K_n(u + k\tau), \, u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right], \, k \in \mathcal{Z}.$$

Kernel-type estimator

- Consider the trigonometric basis in $\mathrm{L}_2[0,\tau]$

$$\phi_1(t) = \sqrt{\frac{1}{\tau}},\ \phi_{2l}(t) = \sqrt{\frac{2}{\tau}}\cos\frac{2\pi l}{\tau}t,\ \phi_{2l+1}(t) = \sqrt{\frac{2}{\tau}}\sin\frac{2\pi l}{\tau}t.$$

Kernel-type estimator

- Consider the trigonometric basis in $\mathrm{L}_2[0,\tau]$

$$\phi_1(t)=\sqrt{\frac{1}{\tau}},\ \phi_{2l}(t)=\sqrt{\frac{2}{\tau}}\cos\frac{2\pi l}{\tau}t,\ \phi_{2l+1}(t)=\sqrt{\frac{2}{\tau}}\sin\frac{2\pi l}{\tau}t.$$

- Coefficients of the kernel-type estimator w.r.t. this basis

$$\tilde{\Lambda}_{1,n}=\hat{\Lambda}_{1,n},\ \tilde{\Lambda}_{2l,n}=\sqrt{\frac{\tau}{2}}K_{2l,n}(\hat{\Lambda}_{2l,n}-\Lambda^*_{2l})+\Lambda^*_{2l},$$

$$\tilde{\Lambda}_{2l+1,n}=\sqrt{\frac{\tau}{2}}K_{2l,n}(\hat{\Lambda}_{2l+1,n}-\Lambda^*_{2l+1})+\Lambda^*_{2l+1},\ l\in\mathcal{N},$$

where $\hat{\Lambda}_{l,n}$ are the Fourier coefficients of the EMF.

Efficiency over a ball

- A kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^\tau K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

Efficiency over a ball

- A kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^\tau K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

- with a kernel satisfying the condition

$$n \sup_{l \geq 1} \left| \sqrt{\frac{\tau}{2}} K_{2l,n} - 1 \right|^2 \longrightarrow 0,$$

Efficiency over a ball

- A kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^\tau K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

- with a kernel satisfying the condition

$$n \sup_{l\geq 1} \left| \sqrt{\frac{\tau}{2}} K_{2l,n} - 1 \right|^2 \longrightarrow 0,$$

- is asymptotically efficient over a ball

$$\lim_{n\to+\infty} \sup_{\Lambda\in\mathcal{B}(R)} \left(\mathbf{E}_\Lambda || \sqrt{n}(\tilde{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) = 0.$$

Efficiency over a compact set

- Now we impose additional conditions of regularity on the unknown mean function

Efficiency over a compact set

- Now we impose additional conditions of regularity on the unknown mean function
- it belongs to $\Sigma(R)$

$$\Sigma(R) = \{\Lambda : \ ||\lambda - \lambda^*||^2 \leq R, \ \Lambda^*(\tau) = \Lambda(\tau)\}.$$

Efficiency over a compact set

- Now we impose additional conditions of regularity on the unknown mean function
- it belongs to $\Sigma(R)$

$$\Sigma(R) = \{\Lambda : \, ||\lambda - \lambda^*||^2 \leq R, \, \Lambda^*(\tau) = \Lambda(\tau)\}.$$

- A kernel-type estimator with the kernel satisfying

$$n \sup_{l \geq 1} \left| \frac{\sqrt{\frac{\tau}{2}} K_{2l,n} - 1}{\frac{2\pi l}{\tau}} \right|^2 \longrightarrow 0,$$

Efficiency over a compact set

- Now we impose additional conditions of regularity on the unknown mean function
- it belongs to $\Sigma(R)$

$$\Sigma(R) = \{\Lambda : \ ||\lambda - \lambda^*||^2 \leq R, \, \Lambda^*(\tau) = \Lambda(\tau)\}.$$

- A kernel-type estimator with the kernel satisfying

$$n \sup_{l \geq 1} \left| \frac{\sqrt{\frac{\tau}{2}} K_{2l,n} - 1}{\frac{2\pi l}{\tau}} \right|^2 \longrightarrow 0,$$

- is asymptotically efficient over $\Sigma(R)$

$$\lim_{n \to +\infty} \sup_{\Lambda \in \Sigma(R)} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\tilde{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t) \mathrm{d}t \right) = 0.$$

Example of another as. effective estimator

- Consider a kernel

$$K(u) \geq 0,\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} K(u)\mathrm{d}u = 1,$$

$$K(u) = K(-u),\quad K(u) = K(u + k\tau),\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ k \in \mathcal{Z}.$$

Example of another as. effective estimator

- Consider a kernel

$$K(u) \geq 0,\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} K(u)\mathrm{d}u = 1,$$

$$K(u) = K(-u),\quad K(u) = K(u + k\tau),\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ k \in \mathcal{Z}.$$

- A sequence $0 \leq h_n \leq 1$ be s.t. $h_n^2 n \longrightarrow 0,\ n \to +\infty.$

Example of another as. effective estimator

- Consider a kernel

$$K(u) \geq 0,\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} K(u)\mathrm{d}u = 1,$$

$$K(u) = K(-u),\quad K(u) = K(u + k\tau),\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ k \in \mathcal{Z}.$$

- A sequence $0 \leq h_n \leq 1$ be s.t. $h_n^2 n \longrightarrow 0,\ n \to +\infty$.
- Then, the kernels

$$K_n(u) = \frac{1}{h_n} K\left(\frac{u}{h_n}\right) \mathbb{1}\left\{|u| \leq \frac{\tau}{2} h_n\right\}$$

satisfy the previous condition and hence

Example of another as. effective estimator

- Consider a kernel

$$K(u) \geq 0,\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} K(u)\mathrm{d}u = 1,$$

$$K(u) = K(-u), \quad K(u) = K(u + k\tau),\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ k \in \mathcal{Z}.$$

- A sequence $0 \leq h_n \leq 1$ be s.t. $h_n^2 n \longrightarrow 0,\ n \to +\infty.$
- Then, the kernels

$$K_n(u) = \frac{1}{h_n} K\left(\frac{u}{h_n}\right) \mathbb{1}\left\{|u| \leq \frac{\tau}{2} h_n\right\}$$

satisfy the previous condition and hence

- the corresponding kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^\tau K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

is as. efficient over $\Sigma(R)$.

Second order efficiency

- How to compare as. efficient estimators?

Second order efficiency

- How to compare as. efficient estimators?
- The first step would be to find the rate of convergence in

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq 0,$$

Second order efficiency

- How to compare as. efficient estimators?
- The first step would be to find the rate of convergence in

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_\Lambda || \sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq 0,$$

- that is, the sequence $\gamma_n \longrightarrow +\infty$ s.t.

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \gamma_n \left(\mathbf{E}_\Lambda || \sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq \mathrm{C},$$

Second order efficiency

- How to compare as. efficient estimators?
- The first step would be to find the rate of convergence in

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq 0,$$

- that is, the sequence $\gamma_n \longrightarrow +\infty$ s.t.

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \gamma_n \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq \mathrm{C},$$

- Then, to construct an estimator which attains this bound.

Second order efficiency

- How to compare as. efficient estimators?
- The first step would be to find the rate of convergence in

$$\underline{\lim}_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq 0,$$

- that is, the sequence $\gamma_n \longrightarrow +\infty$ s.t.

$$\underline{\lim}_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \gamma_n \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq \mathrm{C},$$

- Then, to construct an estimator which attains this bound.
- Calculate the constant C.

Related works

- Second order estimation was introduced by [Golubev G.K. and Levit B.Ya., 1996] in the distribution function estimation problem.

Related works

- Second order estimation was introduced by [Golubev G.K. and Levit B.Ya., 1996] in the distribution function estimation problem.
- For other models second order efficiency was proved by [Dalalyan A.S. and Kutoyants Yu.A., 2004], [Golubev G.K. and Härdle W., 2000].

Related works

- Second order estimation was introduced by [Golubev G.K. and Levit B.Ya., 1996] in the distribution function estimation problem.
- For other models second order efficiency was proved by [Dalalyan A.S. and Kutoyants Yu.A., 2004], [Golubev G.K. and Härdle W., 2000].
- Asymptotic efficiency in non-parametric estimation problems was done for the first time in [Pinsker M.S., 1980],

Related works

- Second order estimation was introduced by [Golubev G.K. and Levit B.Ya., 1996] in the distribution function estimation problem.
- For other models second order efficiency was proved by [Dalalyan A.S. and Kutoyants Yu.A., 2004], [Golubev G.K. and Härdle W., 2000].
- Asymptotic efficiency in non-parametric estimation problems was done for the first time in [Pinsker M.S., 1980],
- where the analogue of the inverse of the Fisher information in non-parametric estimation problem was calculated (Pinsker's constant).

Main theorems

- Introduce

$$\mathcal{F}_m^{per}(R,S) = \left\{ \Lambda(\cdot) : \int_0^\tau [\lambda^{(m-1)}(t)]^2 \mathrm{d}t \leq R,\ \Lambda(0) = 0,\ \Lambda(\tau) = S \right\}$$

where $R > 0$, $S > 0$, $m > 1$, are given constants.

Main theorems

- Introduce

$$\mathcal{F}_m^{per}(R,S) = \left\{ \Lambda(\cdot) : \int_0^\tau [\lambda^{(m-1)}(t)]^2 \mathrm{d}t \leq R,\, \Lambda(0) = 0,\, \Lambda(\tau) = S \right\}$$

where $R > 0$, $S > 0$, $m > 1$, are given constants.

- For all estimators $\bar{\Lambda}_n(t)$ of the mean function $\Lambda(t)$, following lower bound holds

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}_m(R,S)} n^{\frac{1}{2m-1}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq -\Pi,$$

Main theorems

- Introduce

$$\mathcal{F}_m^{per}(R,S)=\left\{\Lambda(\cdot):\int_0^\tau[\lambda^{(m-1)}(t)]^2\mathrm{d}t\leq R,\,\Lambda(0)=0,\,\Lambda(\tau)=S\right\}$$

where $R>0$, $S>0$, $m>1$, are given constants.

- For all estimators $\bar{\Lambda}_n(t)$ of the mean function $\Lambda(t)$, following lower bound holds

$$\varliminf_{n\to+\infty}\sup_{\Lambda\in\mathcal{F}_m(R,S)} n^{\frac{1}{2m-1}}\left(\mathbf{E}_\Lambda\|\sqrt{n}(\bar{\Lambda}_n-\Lambda)\|^2-\int_0^\tau\Lambda(t)\mathrm{d}t\right)\geq-\Pi,$$

- where

$$\Pi=\Pi_m(R,S)=(2m-1)R\left(\frac{S}{\pi R}\frac{m}{(2m-1)(m-1)}\right)^{\frac{2m}{2m-1}},$$

plays the role of the Pinsker's constant.

Second order as. efficient estimator

- Consider

$$\Lambda_n^*(t) = \hat{\Lambda}_{0,n}\phi_0(t) + \sum_{l=1}^{N_n} \tilde{K}_{l,n}\hat{\Lambda}_{l,n}\phi_l(t),$$

where $\{\phi_l\}_{l=0}^{+\infty}$ is the trigonometric cosine basis, $\hat{\Lambda}_{l,n}$ are the Fourier coefficients of the EMF w.r.t. this basis and

$$\tilde{K}_{l,n} = \left(1 - \left|\frac{\pi l}{\tau}\right|^m \alpha_n^*\right)_+, \quad \alpha_n^* = \left[\frac{S}{nR}\frac{\tau}{\pi}\frac{m}{(2m-1)(m-1)}\right]^{\frac{m}{2m-1}},$$

$$N_n = \frac{\tau}{\pi}(\alpha_n^*)^{-\frac{1}{m}} \approx \mathrm{C}n^{\frac{1}{2m-1}}, \quad x_+ = \max(x, 0),\ x \in \mathbf{R}.$$

Second order as. efficient estimator

- Consider

$$\Lambda_n^*(t) = \hat{\Lambda}_{0,n}\phi_0(t) + \sum_{l=1}^{N_n} \tilde{K}_{l,n}\hat{\Lambda}_{l,n}\phi_l(t),$$

where $\{\phi_l\}_{l=0}^{+\infty}$ is the trigonometric cosine basis, $\hat{\Lambda}_{l,n}$ are the Fourier coefficients of the EMF w.r.t. this basis and

$$\tilde{K}_{l,n} = \left(1 - \left|\frac{\pi l}{\tau}\right|^m \alpha_n^*\right)_+, \quad \alpha_n^* = \left[\frac{S}{nR}\frac{\tau}{\pi}\frac{m}{(2m-1)(m-1)}\right]^{\frac{m}{2m-1}},$$

$$N_n = \frac{\tau}{\pi}(\alpha_n^*)^{-\frac{1}{m}} \approx \mathrm{C}n^{\frac{1}{2m-1}}, \quad x_+ = \max(x, 0),\ x \in \mathbf{R}.$$

- The estimator $\Lambda_n^*(t)$ attains the lower bound described above, that is,

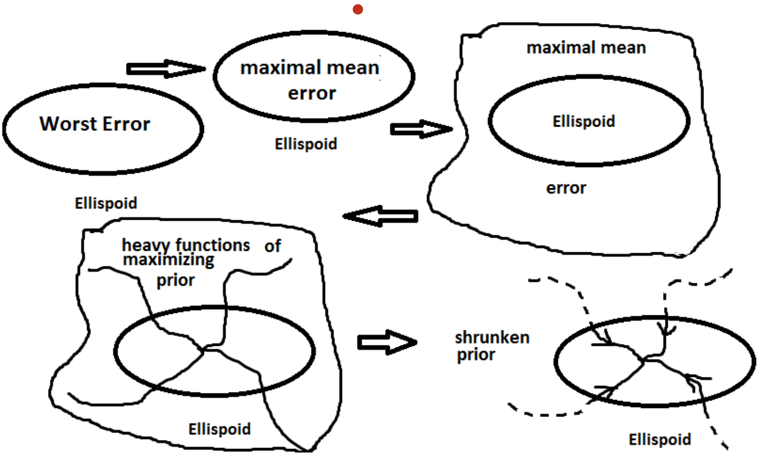
$$\lim_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}_m(R,S)} n^{\frac{1}{2m-1}} \left(\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t\right) = -\Pi.$$

Sketch of the proof

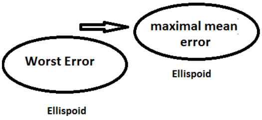
- Proof consists of several steps:

Sketch of the proof

- Proof consists of several steps:

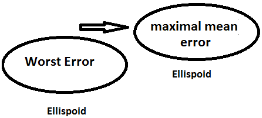


First step



•

First step

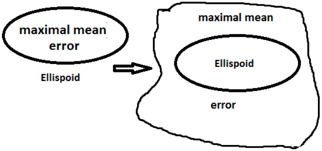


- Reduce the minimax problem to a Bayes risk maximization problem

$$\sup_{\Lambda\in\mathcal{F}_m^{(per)}(R,S)} \left(\mathbf{E}_\Lambda||\bar{\Lambda}_n-\Lambda||^2-\mathbf{E}_\Lambda||\hat{\Lambda}_n-\Lambda||^2\right)\geq$$

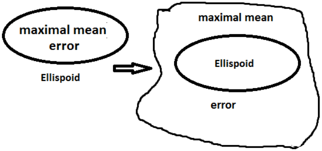
$$\sup_{\mathbf{Q}\in\mathcal{P}}\int_{\mathcal{F}_m^{(per)}(R,S)}\left(\mathbf{E}_\Lambda||\bar{\Lambda}_n-\Lambda||^2-\mathbf{E}_\Lambda||\hat{\Lambda}_n-\Lambda||^2\right)\mathrm{d}\mathbf{Q}.$$

Second step


maximal mean
error
Ellispoid
maximal mean
Ellispoid
error

-

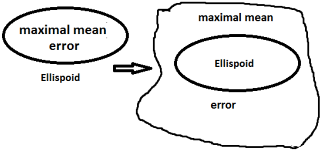
Second step



- In the maximization problem replace the set of probabilities $\mathcal{P}(\mathcal{F})$ concentrated on $\mathcal{F}_m^{(per)}(R,S)$

$$\sup_{\mathbf{Q}\in\mathcal{P}(\mathcal{F})}\int_{\mathcal{F}_m^{(per)}(R,S)}\left(\mathbf{E}_\Lambda||\bar{\Lambda}_n-\Lambda||^2-\mathbf{E}_\Lambda||\hat{\Lambda}_n-\Lambda||^2\right)\mathrm{d}\mathbf{Q},$$

Second step

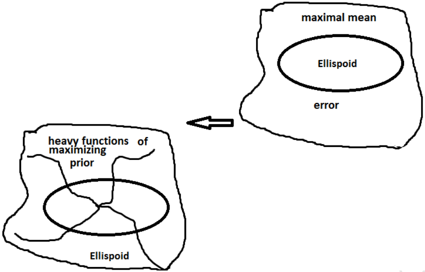


- In the maximization problem replace the set of probabilities $\mathcal{P}(\mathcal{F})$ concentrated on $\mathcal{F}_m^{(per)}(R,S)$

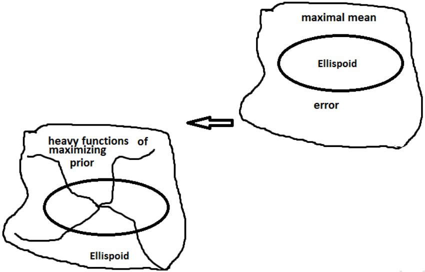
$$\sup_{\mathbf{Q}\in\mathcal{P}(\mathcal{F})} \int_{\mathcal{F}_m^{(per)}(R,S)} \left(\mathbf{E}_\Lambda ||\bar{\Lambda}_n - \Lambda||^2 - \mathbf{E}_\Lambda ||\hat{\Lambda}_n - \Lambda||^2 \right) \mathrm{d}\mathbf{Q},$$

- by the set of probabilities $\mathbf{E}(\mathcal{F})$ concentrated on $\mathcal{F}_m^{(per)}(R,S)$ in mean.

Third step


maximal mean
Ellispoid
error
heavy functions of
maximizing
prior
Ellispoid

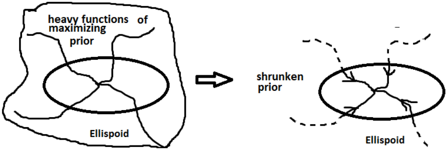
Third step



- Replace the ellipsoid by the least favorable parametric family (heavy functions)

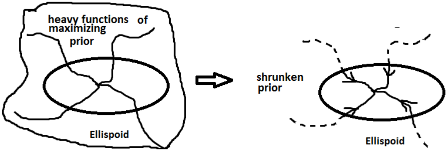
$$\sup_{\Lambda_\theta \in \mathcal{F}_m^{(per)}(R,S)} \int_\Theta \left(\mathbf{E}_\theta ||\bar{\Lambda}_n - \Lambda_\theta||^2 - \mathbf{E}_\theta ||\hat{\Lambda}_n - \Lambda_\theta||^2 \right) \mathrm{d}\mathbf{Q}.$$

Fourth step


heavy functions of maximizing prior
Ellispoid
shrunken prior
Ellispoid

•

Fourth step



- Shrink the heavy functions and the least favorable prior distribution to fit the ellipsoid

$$\mathbf{Q}\{\theta : \Lambda_\theta \notin \mathcal{F}_m^{(per)}(R,S)\} = o(n^{-2}).$$

Further work

- What can be done or what had to be done?

Further work

- What can be done or what had to be done?
- The condition $\Lambda(\tau) = S$ in the definition of the set $\mathcal{F}_m^{(per)}(R, S)$ have to be replaced by $\Lambda(\tau) \leq S$. The last one cannot be thrown out since with a notation $\pi_j(t) = X_j(t) - \Lambda(t)$ we get

$$\hat{\Lambda}_n(t) = \Lambda(t) + \frac{1}{n}\sum_{j=1}^{n} \pi_j(t), \text{ data=signal+“noise”}$$

and the variance of the noise is $\frac{1}{n}\Lambda(t)$. (Simultaneous estimation of the function and its variance).

Further work

- What can be done or what had to be done?
- The condition $\Lambda(\tau) = S$ in the definition of the set $\mathcal{F}_m^{(per)}(R, S)$ have to be replaced by $\Lambda(\tau) \leq S$. The last one cannot be thrown out since with a notation $\pi_j(t) = X_j(t) - \Lambda(t)$ we get

$$\hat{\Lambda}_n(t) = \Lambda(t) + \frac{1}{n}\sum_{j=1}^{n} \pi_j(t), \text{ data=signal+“noise”}$$

  and the variance of the noise is $\frac{1}{n}\Lambda(t)$. (Simultaneous estimation of the function and its variance).
- Adaptive estimation-construct an estimator that does not depend on m, S, R.

Further work

- What can be done or what had to be done?
- The condition $\Lambda(\tau) = S$ in the definition of the set $\mathcal{F}_m^{(per)}(R, S)$ have to be replaced by $\Lambda(\tau) \leq S$. The last one cannot be thrown out since with a notation $\pi_j(t) = X_j(t) - \Lambda(t)$ we get

$$\hat{\Lambda}_n(t) = \Lambda(t) + \frac{1}{n}\sum_{j=1}^{n} \pi_j(t), \quad \text{data=signal+"noise"}$$

  and the variance of the noise is $\frac{1}{n}\Lambda(t)$. (Simultaneous estimation of the function and its variance).
- Adaptive estimation-construct an estimator that does not depend on m, S, R.
- Consider other models or formulate a general result for non-parametric LAN.

References

Gasparyan S.B. and Kutoyants Yu.A.,
On the lower bound in second order asymptotically efficient estimation for Poisson processes.
Submitted. (2016).

Gill R.D. and Levit B.Ya.,
Applications of the van Trees inequality: a Bayesian Cramér-Rao bound.
Bernoulli, 1(1-2), 59-79. (1995).

Golubev G.K. and Härdle W.,
Second order minimax estimation in partial linear models.
Math. Methods Statist., 9(2), 160-175. (2000).

Golubev G.K. and Levit B.Ya.,
On the second order minimax estimation of distribution functions.
Math. Methods Statist., 5(1), 1-31. (1996).

References

Ibragimov, I., Hasminskii R.,
Statistical Estimation: Asymptotic Theory.
Springer, New York (1981).

Dalalyan A.S. and Kutoyants Yu.A.,
On second order minimax estimation of invariant density for ergodic diffusion.
Statistics & Decisions, 22(1), 17-42. (2004).

Kutoyants, Yu. A.,
Introduction to Statistics of Poisson Processes.
To appear. (2016).

Pinsker M.S.,
Optimal filtering of square-integrable signals in Gaussian noise.
Probl. Inf. Transm., 16(2), 120-133. (1980).